

METHOD AND APPARATUS FOR RECOGNITION OF PITCH

Publication number: JP9022298

Publication date: 1997-01-21

Inventor: SZALAY ANDREAS DIPL PHYS (DE)

Applicant: BLUE CHIP MUSIC GMBH (DE); YAMAHA CORP (JP)

Classification:

- international: G10H3/12; G10H3/00; (IPC1-7): G10L9/12; G01R23/02

- european: G10H3/12B

Application number: JP19960021750 19960112

Priority number(s): DE19951000750 19950112

Also published as:



EP0722161 (A2)

US5780759 (A1)

EP0722161 (A3)

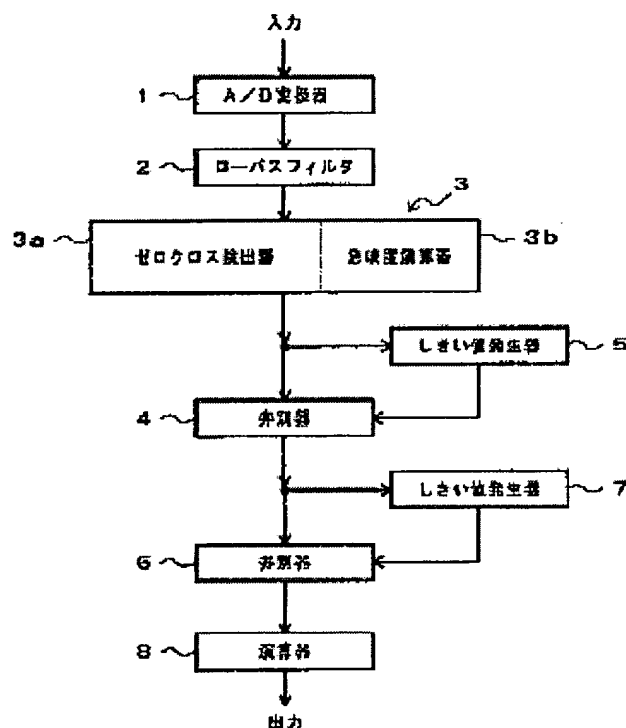
DE19500750 (A1)

EP0722161 (B1)

Report a data error here

Abstract of JP9022298

PROBLEM TO BE SOLVED: To simply and surely recognize a sound pitch of a high level signal generated from a musical instrument, etc., by detecting a pitch of a signal waveform in each zero-crossing area and using the detected pitch as an evaluation reference to select a zero cross to be evaluated. **SOLUTION:** A pitch of a signal waveform is detected in each zero-crossing area and the detected pitch is used as an evaluation reference to select a zero cross to be evaluated. An inputted signal waveform is filtered by a low-pass filter 2 through an A/D converter 1. The output of the low-pass filter 2 is inputted to a computing unit 3 constituted of a zero-cross detector 3a and a steepness computing element 3b. The zero cross detector 3a in the computing unit 3 detects the zero cross of the output of the low-pass filter 2 and the timing of the zero cross. The steepness computing unit 3b calculates steepness in the zero cross waveform at every zero crossing.



Data supplied from the esp@cenet database - Worldwide

<paragraph 28>

At this point, by setting

5 $D10 = a1 - 10$

$$D21 = A2 - A1$$

$$D32 = A3 - A2$$

$Dx = A2 / (A2 - A1)$ (distance between the zero-crossing point and the time point P2)

10 then slope value D is

$$D = (D21 + dx \times D10 + (1 - dx) \times D32) / 2.$$

When it is desired to avoid floating-point arithmetic, it is acceptable to interpolate using integer calculation when simulating "oversampling" 16 times. When the ratio between individual slope value is considered instead of absolute slope, one can avoid division by 2. In this case, the following can be set:

15

$$dx = (a2 \ll 4) / (A2 - A1)$$

$$D = \{dx \times (A2 - A0) + (16 - dx) \times (A3 - A1)\}$$

Here, the symbol "<<" indicates "left shift" of the binary number. Shifting 4

20 bits to the left is equal to multiplying by 16, and at the time point of the zero-crossing,

$$T = (1X \ll 4) - dx$$

Here, 1X is the sampling index number at the time point P2. The length of period is obtained from the difference in the consecutive zero-crossing

25 points detected as shown.

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-22298

(43) 公開日 平成9年(1997)1月21日

(51) Int.Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 9/12			G 1 0 L 9/12	B
G 0 1 R 23/02			G 0 1 R 23/02	

審査請求 未請求 請求項の数18 F D (全 10 頁)

(21) 出願番号 特願平8-21750
(22) 出願日 平成8年(1996)1月12日
(31) 優先権主張番号 1 9 5 0 0 7 5 0 . 6 - 5 1
(32) 優先日 1995年1月12日
(33) 優先権主張国 ドイツ (D E)

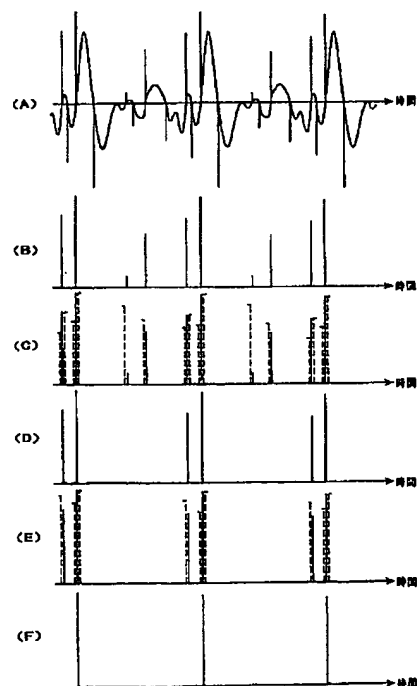
(71) 出願人 596017565
ブルー チップ ミュージック ジーエム
ピーエイチ
ドイツ国 D-56283 ハルゼンバッハ
インダストリエストラッセ 3
(71) 出願人 000004075
ヤマハ株式会社
静岡県浜松市中沢町10番1号
(72) 発明者 アンドレアス スツァレー
ドイツ国 D-56281 エメルシャウゼン
リンデンストラッセ 1
(74) 代理人 弁理士 飯塚 義仁

(54) 【発明の名称】 音高認識方法及び装置

(57) 【要約】

【課題】 例えば爪弾き操作または打ち鳴らし操作によって励振される楽器などから発生された音信号の音高認識（ピッチ検出）を簡単かつ確実に可能にする。

【解決手段】 入力された音信号の信号波形のゼロクロス間の間隔を該音信号の周期長さ測定値として使用することによって音高を認識する方法または装置において、前記信号波形の傾斜の大きさを、ゼロクロスの領域のそれぞれについて検出し、検出した前記傾斜の大きさを評価規準として使用して、評価すべきゼロクロス（すなわち音高認識に使用するゼロクロス）を選択する。



【特許請求の範囲】

【請求項1】 入力された音信号の信号波形のゼロクロス間の間隔を該音信号の周期長さ測定値として使用することによって音高を認識する音高認識方法において、前記信号波形の傾斜の大きさをゼロクロスの領域のそれぞれについて検出し、検出した前記傾斜の大きさを、評価すべきゼロクロスを選択するための評価規準として使用することを特徴とする音高認識方法。

【請求項2】 入力された音信号の信号波形のゼロクロス間の間隔を該音信号の周期長さ測定値として使用することによって音高を認識する音高認識方法において、前記信号波形の傾斜の大きさをゼロクロスの領域のそれぞれについて検出するステップと、検出した前記信号波形の傾斜の最大値を判定し、該傾斜の最大値に基づいて減衰関数を発生し、現時点における前記減衰関数の値を超える大きさの傾斜を有するゼロクロスについてのみ以後の処理を行い、こうして評価すべきゼロクロスを選択するステップとを備えた音高認識方法。

【請求項3】 ゼロクロスが発生した時のみ、前記減衰関数の値を小さくする請求項2に記載の音高認識方法。

【請求項4】 前記減衰関数の値を小さくしようとする毎に、該減衰関数の値に一定の係数を乗ずる請求項3に記載の音高認識方法。

【請求項5】 前記減衰関数の値を超える傾斜値を、同じやり方で少なくとももう一度、前記減衰関数と比較する請求項2に記載の音高認識方法。

【請求項6】 ゼロクロスにおける前記傾斜を、該ゼロクロスの近辺における前記音信号の複数の傾斜値によって補間する請求項1に記載の音高認識方法。

【請求項7】 ある特定のゼロクロスにおける前記傾斜が次のゼロクロスの傾斜の大きさの所定比率に達しない場合、前記特定のゼロクロスを除外する請求項1に記載の音高認識方法。

【請求項8】 ゼロクロスの間の連続した時間間隔を相互比較し、前記時間間隔の間の差が所定のリミット未満である場合にのみ、音高を決定する請求項1に記載の音高認識方法。

【請求項9】 前記音信号について固定したサンプリング周波数を使用し、前記時間間隔において決定された音高を平均化することによって、所定の長さを有する時間間隔の終わりにのみ、音高オリジナル値を発生する請求項1に記載の音高認識方法。

【請求項10】 前記音高オリジナル値が直ぐ前に通過させられた音高オリジナル値から所定の量だけ異なる場合にのみ、インターフェイスを介して前記音高オリジナル値を通過させる請求項9に記載の音高認識方法。

【請求項11】 音高決定の前に、前記音信号をローパ

スフィルタ処理する請求項1に記載の音高認識方法。

【請求項12】 ゼロクロスを正方向および負方向の両方向に評価する請求項1に記載の音高認識方法。

【請求項13】 ある特定のゼロクロスの傾斜が、直ぐ前のゼロクロスの逆極性の傾斜の半分未満である場合、前記特定のゼロクロスを評価しないようにする請求項12に記載の音高認識方法。

【請求項14】 入力された音信号の信号波形のゼロクロス間の間隔を該音信号の周期長さ測定値として使用することによって音高を認識する音高認識方法において、前記信号波形の傾斜の大きさをゼロクロスの領域のそれぞれについて検出する処理と、検出した前記傾斜の大きさを、評価すべきゼロクロスを選択するための評価規準として使用する処理と、重要なゼロクロスの時点を補間によって決定する処理とを備えた音高認識方法。

【請求項15】 爪弾きまたは打ち鳴らし操作によって励振される楽器から発生された楽音信号を前記音信号として入力する請求項1乃至14のいずれかに記載の音高認識方法。

【請求項16】 時間関数としての振幅値 $A(t)$ からなる波形によって表わされる楽音の音高を決定するための音高認識装置であって、前記波形が前記音高を定義する略等しい長さの複数の周期を有し、前記波形の各周期が $A(t) = 0$ である複数のゼロクロスを有し、(a) 前記波形の少なくとも1つの周期における前記波形のゼロクロスを検出するゼロクロス検出手段と、(b) 各前記ゼロクロスごとに前記波形の急峻度値を求める急峻度算出手段と、(c) しきい値を発生するしきい値発生手段と、(d) 前記急峻度算出手段によって求められた急峻度値を前記しきい値と比較することによって、前記検出されたゼロクロスのうち前記しきい値未満の急峻度値を有するゼロクロスを弁別し、これにより、前記少なくとも1つの周期における最終的なゼロクロスを決定する弁別手段と、(e) 前記少なくとも1つの周期の長さを定義する前記最終的なゼロクロスに基づいて、前記音高を算出する演算手段とを具備する音高認識装置。

【請求項17】 時間関数としての振幅値 $A(t)$ からなる波形によって表わされる楽音の音高を決定するための音高認識装置であって、前記波形が前記音高を定義する略等しい長さの複数の周期を有し、前記波形の各周期が $A(t) = 0$ である複数のゼロクロスを有し、(a) 前記波形の少なくとも1つの周期における前記波形のゼロクロスを検出するゼロクロス検出手段と、(b) 各前記ゼロクロスごとに前記波形の急峻度値を求める急峻度算出手段と、(c) ゼロクロスが発生するごとに変更される動的なしきい値を発生するしきい値発生手段と、(d) 前記急峻度算出手段によって求められた急峻度値を前記しきい値と比較することによって、前記検出されたゼロクロスのうち前記しきい値未満の急峻度値を有す

るゼロクロスを弁別し、これにより、前記少なくとも1つの周期における最終的なゼロクロスを決する弁別手段と、(e)前記少なくとも1つの周期の長さを定義する前記最終的なゼロクロスに基づいて、前記音高を算出する演算手段とを具備する音高認識装置。

【請求項18】 前記動的なしきい値を、前記しきい値を超える急峻度値を有するゼロクロスが発生するごとに大きくなり、次のゼロクロスの急峻度値と比較する前毎に小さくなるよう、変更する請求項17に記載の音高認識装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は、例えば爪弾き操作または打ち鳴らし操作によって励振される楽器のような適宜の音発生源から発生された音信号の音高(すなわちピッチ)を認識する方法及び装置に関し、特に、該音信号の信号波形のゼロクロス間の間隔を該音信号の周期長さ測定値として使用することによって該音高を認識する方法及び装置に関する。

【0002】

【従来の技術】合成音発生技術が利用され始めた初期には、各キーに対して明確に定義された音が割り当てられた鍵盤楽器が基準楽器として使用されていたが、長い間、合成音を発生するために他の種類の楽器を使用する試みがなされてきた。このような楽器の一例としては、指またはピックを使用した爪弾き操作または打ち鳴らし操作によって、張設された弦が振動されるギターがある。既知のように、ギターの場合、その弦の有効振動長さを変えることによって、様々な音高を発生できる。伝統的なアコースティックギターの場合、弦の振動はギター本体の共鳴によって直接的に耳に聞こえる楽音を発生するが、電子的な楽音合成技術を適用しようとする場合には、励振された弦の振動周波数(すなわち音高若しくはピッチ)を検出する必要がある。音高が検出されれば、その音高情報に対応する楽音信号を適宜の楽音合成技術を利用して合成することができ、さらに適宜に処理することができる。ギターのみならず、例えばハープ、ベース、チターなどの爪弾き操作または打ち鳴らし操作される他の弦楽器についても同様の課題があるであろう。また、ドラムの場合にも、音高認識が重要になることがある。更に、原則的には、このような音高認識方法は、いわゆる“ボイスフォロア(voice follower)”でさらに処理可能な人間の音声における音高認識にとっても必要であり、その他すべての音信号(すなわちオーディオ信号)に適用可能である。しかし、便宜上、ここではギターにおける音高認識について以下に説明する。

【0003】米国特許第5,014,589号には、音信号のゼロクロスを検出する音高認識方法が開示されている。この特許によると、同一方向の2つのゼロクロス間の間隔が音信号の周期長さ測定値として考慮される。この周期

の逆数が周波数に相当する。

【発明が解決しようとする課題】しかし、上記特許に開示されたような音高認識方法の問題点は、周期を決するゼロクロスに加えて、例えば高調波に起因するゼロクロスも1つの周期内に発生する、ということである。従って、上記音高認識方法にあつては、ゼロクロスの時点のみならず、信号波形の振幅最大値をも検出する必要がある。この場合、“エンベロープフォロア(envelope follower)”とも呼ばれる一種のエンベロープ曲線が発生される。その結果、ある特定のゼロクロスが1つの周期の境目を表わしているか否かを判定するために、付加的な規準が使用されている。2つの連続した周期が所定量以上異なっていないときに、音高信号が発生される。

【0004】上記のような方法における信号処理は、ますます、デジタル的に行われるようになっていく。公知の音高認識方法にあつては、計算のために相当量のエネルギーが必要である。1つの弦だけでなく複数の弦について継続的に計算のためのエネルギーを確保しなければならぬので、今日利用可能なプロセッサによっては、経済的な解決を実現することは実質的に不可能である。この発明は上述の点に鑑みてなされたもので、簡単に、確実な音高認識をおこなうことができる方法及び装置を提供することを目的とする。

【0005】

【課題を解決するための手段】上記目的を達成するために、この発明に係る方法は、入力された音信号の信号波形のゼロクロス間の間隔を該音信号の周期長さ測定値として使用することによって音高を認識する音高認識方法において、前記信号波形の傾斜の大きさをゼロクロスの領域のそれぞれについて検出し、検出した前記傾斜の大きさを、評価すべきゼロクロスを選択するための評価規準として使用することを特徴とするものである。

【0006】また、この発明に係る方法は、入力された音信号の信号波形のゼロクロス間の間隔を該音信号の周期長さ測定値として使用することによって音高を認識する音高認識方法において、前記信号波形の傾斜の大きさをゼロクロスの領域のそれぞれについて検出するステップと、検出した前記信号波形の傾斜の最大値を判定し、該傾斜の最大値に基づいて減衰関数を発生し、現時点における前記減衰関数の値を超える大きさの傾斜を有するゼロクロスについてのみ以後の処理を行い、こうして評価すべきゼロクロスを選択するステップとを備えたことを特徴とするものである。更に、この発明に係る方法は、入力された音信号の信号波形のゼロクロス間の間隔を該音信号の周期長さ測定値として使用することによって音高を認識する音高認識方法において、前記信号波形の傾斜の大きさをゼロクロスの領域のそれぞれについて検出する処理と、検出した前記傾斜の大きさを、評価すべきゼロクロスを選択するための評価規準として使用する処理と、重要なゼロクロスの時点を補間によって決定

する処理とを備えたことを特徴とするものである。

【0007】また、この発明に係る装置は、時間関数としての振幅値 $A(t)$ からなる波形によって表わされる楽音の音高を決定するための音高認識装置であって、前記波形が前記音高を定義する略等しい長さの複数の周期を有し、前記波形の各周期が $A(t)=0$ である複数のゼロクロスとを有し、(a)前記波形の少なくとも1つの周期における前記波形のゼロクロスを検出するゼロクロス検出手段と、(b)各前記ゼロクロスごとに前記波形の急峻度値を求める急峻度算出手段と、(c)しきい値を発生するしきい値発生手段と、(d)前記急峻度算出手段によって求められた急峻度値を前記しきい値と比較することによって、前記検出されたゼロクロスのうち前記しきい値未満の急峻度値を有するゼロクロスを弁別し、これにより、前記少なくとも1つの周期における最終的なゼロクロスを決する弁別手段と、(e)前記少なくとも1つの周期の長さを定義する前記最終的なゼロクロスに基づいて、前記音高を算出する演算手段とを具備したものである。更に、この発明に係る装置は、上記しきい値発生手段が、ゼロクロスが発生するごとに変更される動的なしきい値を発生することを特徴とする。

【0008】この発明に係る上記方法によると、上記目的は、前記信号波形の傾斜の大きさをゼロクロスの領域のそれぞれについて検出し、検出した前記傾斜の大きさを、評価すべきゼロクロスを選択するための評価規準として使用することによって達成される。必要とされる計算エネルギーは、米国特許第5,014,589号に開示された方法に比べて約 $1/10$ 以下に大幅に減少可能である。つまり、サンプル値をディジタル変換した信号としての音信号は、ゼロクロス領域についてのみ評価される。ゼロクロスは、2つの連続したサンプル値の極性を比較することによって容易に検出可能である。その他のサンプル値はすべて、評価の対象から除外される。精度を高めるために必要な場合、ゼロクロス領域内の数個の値が付加的に考慮される。同様に、ゼロクロスの傾斜も比較的容易に検出可能である。一定のサンプリング周波数を前提とした場合、原則的に、ゼロクロスの前、後の2つのサンプル値の間の間隔を検出すればよい。音信号の波形は、周期を規定するゼロクロスにおいて最も急峻になるからである。従って、考慮すべきことは、同じ極性の最も急峻なゼロクロスだけである。こうして、これらのゼロクロスの間の間隔が、周期長さとして決定される。このようにして、ある特定のゼロクロスが周期長さを求めるために重要であるか否かを判定するために必要な情報は、ゼロクロスにおける信号波形から直接的に得られる。このように、ゼロクロスまたはその直ぐ近辺に位置するサンプル値のみを計算対象とすればよいので、必要な計算エネルギーを大幅に減少させることができる。

【0009】さらに、信号波形が最も急峻な、すなわち、信号波形の傾斜が最も大きなゼロクロスを使用する

ことによって、信号波形の乱れによる影響を最小にできる、という利点がある。最も簡単な構成例において、このような乱れをオフセット（すなわち、信号波形の正方向または負方向におけるズレ）とみなした場合、平坦な信号波形がゼロ軸に交差する際のズレは、急峻な信号波形のゼロクロスより大きくなる。このような急峻なゼロクロスに制限することによって、音高認識の精度が高められる。

【0010】ゼロクロス周辺の比較的狭い帯域以外の音信号波形についての情報は必要とされないで、比較的粗い分解能、すなわち、低いサンプリングレートによる適切な処理が可能である。人間の耳は、自己の周波数帯域において、比較的細かな分解能を有する。このため、音高情報は、約1セント、すなわち、半音の $1/100$ の精度で実現されなければならない。80Hz～1kHzの周波数帯域を有するギターの場合、このような目的のためには、1.7MHzのサンプリングレートが必要であり、必要な計算の複雑さは莫大である。この発明に係る方法を使用することによって、従来の方法に比べてはるかに少ない数のサンプル値だけで、適切な処理を行うことができる。この場合、約10kHzのサンプリングレートが適当である。

【0011】評価に使用すべき信号波形の傾斜を分別するために、好ましくは前記傾斜の最大値が検出され、この最大値に基づいて減衰関数が発生され、現時点における減衰関数の値を超える傾斜値を有するゼロクロスのみが、以後の処理を受ける。一方、前記減衰関数によって、小さすぎる傾斜を有するすべてのゼロクロスが除外される。さらに、前記以後の処理の間、これらのゼロクロスのための計算エネルギーは必要でない。このようにして、重要ではないゼロクロスの除外は比較的早期に行われる。さらに、固定したしきい値とは対照的に、前記減衰関数は、実際の楽器のダイナミックレンジが考慮される、という利点を有する。また、前記信号波形の傾斜は、とりわけ、楽器の演奏音量によって決定される。さらに、弦が打ち鳴らし操作された瞬間には、前記傾斜にスパイクが発生することがある。このようなスパイクは基本的に重要な（意義ある）ものではない。前記減衰関数は、楽器のダイナミックレンジとの整合にも関わらず、小さすぎる傾斜を有するゼロクロスの除外を可能にするが、上述したスパイクがこの音高認識方法の実施を妨害しないことを保証する。

【0012】本発明によると、ゼロクロスが発生した時にのみ前記減衰関数の値を小さくすることが特に好ましい。このようにして、計算エネルギーを節約でき、一方、前記減衰関数を段階的に小さくすることもできる。さらに、前記減衰関数を小さくしようとする毎に、前記減衰関数の値に一定の係数を乗じることも好ましい。このようにして、初期的には大幅な値の減少を行い、その後は、穏やかな値の減少を行うような、指数的な減衰が実

現される。従って、スパイクはより速く除去される。

【0013】重要なものとして残された傾斜値は、同じやり方で少なくとももう一度、減衰関数と比較されるのが好ましい。こりようにして、評価能力を高めることができる。特に打ち鳴らし操作によるスタート領域における音信号の自然な不均等性の結果、前記傾斜値に比較的大きな分散が生じる。前記しきい値が大きすぎ値である場合、認識されるべき重要なゼロクロスが認識され得ない。音信号が多数のゼロクロスをもつ場合、前記減衰関数は高速で小さすぎる値に達するので、前記減衰関数と前記傾斜との比較のによって、重要ではないゼロクロスが間違っって“重要”と分類されることになる。二次的（付加的）な“フィルタ処理”は、依然として正しくない、すなわち、不必要な値を除外するが、重要な値はすべて確実に保持する。原則として、周期長さの決定に使用される最も急峻なゼロクロスを実際に検出するためには、1回の二次的な比較で十分である。

【0014】好ましくは、ゼロクロスにおける前記傾斜は、該ゼロクロスの近辺の音信号の複数の傾斜値によって補間される。ベースがゼロクロス領域における基本的にリニアな信号波形である場合には、2つの値に基づく1回の傾斜検出で十分である。しかし、この領域における信号波形が比較的急な曲率を有するものである場合、このような簡単な傾斜検出ではエラーが発生する。

【0015】好ましくは、ある特定のゼロクロスにおける前記傾斜が次のゼロクロスの傾斜の大きさの所定比率に達しない場合、前記特定のゼロクロスは重要ではないものとして除外される。このようにして、スパイク、すなわち、通常の信号波形に適合しない値が、容易且つ速やかに除外され得る。好ましくは、重要なゼロクロスの時点は、補間によって決定される。しかし、このような補間は、重要なゼロクロスが検出された場合にのみ必要である。このため、有用な結果が実際に予期されているときにのみ、計算エネルギーが必要になる。

【0016】好ましくは、ゼロクロスの間の連続した時間間隔が相互比較し、所定のリミット未満の差がある場合にのみ、音高が決定される。これは、音高およびこれに関連した周期長さがテーブルに格納される場合、特に有利である。前記周期長さが変化しない限り、音高は変化しない。従って、必要な情報は既に存在しているので、情報を得るための新たな計算またはサーチは不必要である。これも、相当な計算時間の節約になる。

【0017】さらに好ましい実施の態様において、前記音信号について固定したサンプリング周波数が使用され、前記時間間隔において決定された音高が平均化されることによって、所定の長さを有する時間間隔の終わりにのみ、音高オリジナル値が発生される。前記時間間隔は、例えば8～15msの長さを有するものである。固定したサンプリング周波数では、音が低くなればなるほど周期当たりのサンプル値が多くなの、音が高くなれば

なるほど周期当たりのサンプル値が少なくなる。従って、本質的に、音が高くなればなるほど音高認識精度が低くなる。この不都合は、固定した時間間隔内に平均化することによって補償される。個々の周期における相対的な精度は、明らかに幾分低いものになる。しかし、高い音の場合については、固定した時間間隔内により多くの周期が含まれるという事実によって、実際の音高により忠実に近づくよう再度の平均化が行われる。

【0018】特に好ましくは、前記音高オリジナル値が直ぐ前に通過させられた音高オリジナル値から所定の量だけ異なる場合にのみ、インターフェイスを介して前記音高オリジナル値が通過させられる。このようなインターフェイスは、例えば、MIDI (Musical Instrument Digital Interface) であってよく、他の形態の信号伝送に依然として広く使用されているものである。伝送されるデータの変化を制限することによって、前記インターフェイスを空き状態にし続けることができる。

【0019】前記音信号は、音高認識の前に、ローパスフィルタ処理されるのが好ましい。このようなローパスフィルタ処理は、例えば2極IIRフィルタを使用することによって、極めて情報を過剰に除外することを回避するよう極めて慎重に行う必要がある。参考的な数字として、このフィルタ処理の後、10を超えるゼロクロスが残らないようにしてよい。

【0020】好ましく、ゼロクロスは、正方向および負方向の両方に評価される。明らかに、この評価法は、1つの極性だけの評価する場合より、多くの計算エネルギーを必要とする。しかし、精度向上に貢献する付加的な情報が得られる。特に好ましくは、ある特定のゼロクロスの傾斜が直ぐ前のゼロクロスの逆極性の傾斜の半分未満である場合、前記特定のゼロクロスは評価されない。従って、この場合、周期長さを決定するためゼロクロスの使用は行われない。しかし、他方の極性のゼロクロスの間の間隔から周期長さを決定できるので、この情報ロスに対処できる。

【0021】

【発明の実施の形態】以下、添付図面を参照してこの発明の実施の形態を詳細に説明しよう。図1は、各周期Tごとに複数のゼロクロスが存在する典型的な音信号（オーディオ信号）の波形例を示す図。ここに例示された音信号は、簡単な2極IIRフィルタを使用した低域通過フィルタ処理が既になされたものである。前記フィルタは乱れを起こす高調波を除去する。この音信号は、以後の処理のためにデジタル化される。すなわち、様々な時点P0, P1, P2, P3, ...における振幅値A0, A1, A2, A3, ...が検出され（図3）、デジタル値に変換される。これらのデジタル値は、3つ以上の値を格納するシフトレジスタまたはFIFOバッファに格納されることである。

【0022】図1に示した信号波形のゼロクロスは、2

つの連続したサンプル値を相互比較することによって容易に検出可能である。つまり、例えばA0とA1、A2とA3のように2つの連続したサンプル値が同一の極性を有する場合、ゼロクロスは存在しない。ゼロクロスの直ぐ近辺における例外を無視する場合、このような値は除外することができる。周期長さTは、2つのゼロクロス間のインターバル、すなわち、例えば、X21P-X11P、X22P-X12P、X21N-X11NまたはX22N-X12N間のインターバルである。このような周期長さの検出は任意の方法で行ってもよいが、X21PとX11PまたはX21NとX11Nの1対の値を使用する場合、最も正確な結果が得られる。というのは、前記信号波形は、これらの時点のゼロクロスにおいて最も大きな傾斜度を有し、ここでは、波形の乱れによる影響が最も小さいからである。すなわち、ゼロクロスのオフセットが小さければ小さいほど、該ゼロクロスにおける信号波形の傾斜が急になる。

【0023】図2を参照して、最も急峻なゼロクロスの検出に使用される比較的簡単な方法について説明する。図2の(A)は、各周期ごとに複数のゼロクロスをもつ典型的な信号波形を示し、さらに、各ゼロクロスにおける信号波形の傾斜の大きさを示す図。図2の(B)は正の傾斜値を示す。各図示例において、これらの傾斜値は、それぞれのゼロクロスに隣接した2つのサンプル値を減算することによって、簡単に決定される。この実施の形態におけるサンプリングレートは10kHzの一定レートであるので、前記減算によって得られる差は、前記傾斜についてのステートメントを作成するのに十分なものになる。

【0024】図2の(A)と(B)とを比較することによって、それ以上の評価のために大量の情報が必要でなくなる、ということが理解されるであろう。こうして、大量の情報のための計算エネルギーが必要でなくなる。図2の(C)は、図2の(B)の傾斜値を示すと共に、破線によって減衰関数の値を示す。該減衰関数は次のようにして形成される。ここで、Dは傾斜値、ENV1は減衰関数の値、F1は例えば11/16である一定の減衰係数を示す。

【0025】一番目のゼロクロスにおいて、ENV1が値Dに設定される。二番目のゼロクロスにおいて、 $ENV1 = F1 \times ENV1$

によって表わされるように、前記減衰関数が変更される。ここで、 $D > ENV1$ ならば、 $ENV1 = D$ が設定される。この例は、二番目のゼロクロスについて示されている。もし、 $D < ENV1$ ならば、これは、重要ではないものと考えられる小さな傾斜を有するゼロクロスである。従って、この箇所は以後の評価から除外される。

【0026】図2の(D)から理解されるように、このような最初のフィルタ処理の後、一番目、二番目、五番

目、六番目、九番目、十番目等のゼロクロスのみが残り、その他のゼロクロスはすべて除去された状態になる。重要なものとして残された前記ゼロクロスは、上記と同様に、

$$ENV2 = F2 \times ENV2$$

によって表わされるように、次のフィルタ処理(図2の(E))を受ける。ここで、ENV2は減衰関数の値、F2は減衰係数を示す。D > ENV2の場合にのみ、対応するゼロクロスはさらに評価され、D > ENV2ではない場合、対応するゼロクロスは重要ではないものとして除外される。

【0027】図2の(F)から理解されるように、図2の(E)の後には、最も急峻なゼロクロスのみが最終的に残る。これらのゼロクロスの間の間隔は周期長さTであり、該周期長さTが音高の測定値である。音高測定精度を高めるために、ゼロクロスの近傍の時点を使用してもよい。例えば、隣接した2つの時点P1、P2そのものではなく、これらの前後の時点P0、P3を使用してもよい。

【0028】ここで、

$$D10 = A1 - A0$$

$$D21 = A2 - A1$$

$$D32 = A3 - A2$$

$$dx = A2 / (A2 - A1) \quad (\text{ゼロクロスと時点P2との間の距離})$$

とすると、傾斜値Dは、

$$D = (D21 + dx \times D10 + (1 - dx) \times D32) / 2$$

となる。浮動小数点演算を回避したい場合において、16回の“オーバーサンプリング”をシミュレートするとき、整数演算を使用した補間を行ってもよい。絶対的な傾斜に注目せず、個々の傾斜値の間の比に注目する場合、2による除算を回避できる。この場合、

$$dx = (A2 \ll 4) / (A2 - A1)$$

$$D = \{ dx \times (A2 - A0) + (16 - dx) \times (A3 - A1) \}$$

を設定することができる。ここで、記号“<<”は、二進数の“左シフト”演算を示す。左方向に4ビットシフトすることは16を掛けることに等しく、ゼロクロスの時点は、

$$T = (1X \ll 4) - dx$$

となる。ここで、1Xは時点P2におけるサンプリング指数である。このようにして検出された2つの連続したゼロクロス時点の差によって、周期長さが得られる。

【0029】2つの連続した周期長さの差が所定の値(例えば40~60セント)未満である場合、上記のように検出された周期長さは、実際に、弦の振動周期の長さに対応する、ものであるとみなすことができる。この場合、小さな誤差を除去するために、前記周期長さは、2つの連続した周期長さの平均値とされる。

【0030】他の誤差修正方法は、連続した値を後方向に相互比較することによって提供される。例えば、傾斜値50, 35, 27のシーケンスが考えられる。これは、高速で減衰している信号に対応する。一方、傾斜値50, 35, 48のシーケンスは、比較的あり得ないものである。この場合、二番目の傾斜値“35”はその信号に適當ではなく、従って、これに関連するゼロクロスを除外する必要がある。これは、前の値を現在値の所定比率と比較することによって、比較的容易に実現可能である。F3が例えば3/4である1より小さい一定値であるとした場合において、

$$F3 \times D(n) > D(n-1)$$

のとき、傾斜値 $D(n-1)$ に関連するゼロクロスが除外される。

【0031】ここに説明した方法の絶対的な精度は $\pm 1/32T$ である。なお、ここで、 T はサンプリング周期である。また、相対的な精度は周波数に左右される。つまり、低い周波数にあっては相対的な精度が高くなり、従って、1セント（半音の $1/100$ ）の誤差の信号を発生するのに十分である。しかし、周波数が高くなるのにつれて相対的な誤差が大きくなり、正しくない音高情報が発生するおそれがある。この誤差は、例えば8~15msの一定長さの“タイムスロット”の終わり以外の各周期の終わりに、音高信号の発生を停止することによって克服される。いずれにせよ、後続の処理は対応する時間を取るのので、音高情報を高速で供給する必要はない。このようなタイムスロットにあっては、周波数が低ければ低いほど、高い相対的な精度で少ない周期が検出される。逆に、周波数が高ければ高いほど、低い相対的な精度で多くの周期が検出される。各々のタイムスロットにおける周期長さが平均化される場合、人間の耳にとって不快でなくなる程度まで誤差を解決できる。

【0032】周期長さ、つまり、音高情報は、正の傾斜を持つゼロクロスおよび負の傾斜を持つゼロクロスの両方から得られる。これらの傾斜の大きさが互いに著しく異なる情況が、時折起こる。一方の傾斜量が他方の傾斜量に比べて3倍以上大きい場合、傾斜量が小さい方のゼロクロスは考慮されない。音高決定中にゼロクロスが評価されるようにするためには存在しなければならない最小の傾斜を定義することもできる。この最小の傾斜は、前のタイムスロットの最大傾斜の $1/2$ を次のタイムスロットの最小傾斜として使用することによって、動的に変化可能である。

【0033】図4は、この発明に係る音高認識装置を示すブロック図である。ギター等の弦楽器のピックアップから入力された波形信号は、音声入力信号としてA/D変換器1に与えられ、該A/D変換器1において、一定のサンプリングレートでサンプリングされ、デジタル信号に変換される。該デジタル信号は、乱れを起こす高調波を除去するために、ローパスフィルタ2でフィル

タ処理される。図2の(A)に示すような波形を有する該ローパスフィルタ2の出力は、ゼロクロス検出器3aと急峻度演算器3bとで構成される演算ユニット3に与えられる。該ユニット3において、前記ゼロクロス検出器3aは、前記ローパスフィルタ2の出力のゼロクロスを検出し、上述した方法のうちの1つに従って、ゼロクロスのタイミングを検出する。前記急峻度演算器3bは、各ゼロクロスごとに、該ゼロクロスの波形における急峻度を算出する。該急峻度を算出するためのいくつかの方法は、上述されている。前記急峻度を算出する最も簡単な方法は、対応するゼロクロスの直ぐ近辺における2つのサンプル値の絶対値を算出する方法である。

【0034】前記ゼロクロス検出器3aおよび急峻度演算器3bは、前記ローパスフィルタ2から与えられるデータの量を大幅に減少させる。前記演算ユニット3の出力は一例のデータ対からなり、各データ対の一番目のデータはゼロクロスのタイミング位置を示し、二番目のデータは対応するゼロクロスの波形の急峻度を示す。比較的小さな急峻度を有するゼロクロスを除外するために、前記演算ユニット3の出力は弁別器4に与えられる。該弁別器4は、所定のしきい値より小さな急峻度を有するゼロクロスのすべてを除外するためのものである。しきい値発生器5は、上述した方法に従って、しきい値ENV1を発生する。短く言えば、前記しきい値ENV1は、各ゼロクロスごとに一定係数F1ずつ小さくされ、また、急峻度値が前のしきい値より大きくなった場合にそのゼロクロスの急峻度値にまで大きくされる。

【0035】データ量が図2の(D)に例示されたデータの量に減少するよう、前記弁別器4は、比較的小さな急峻度を有するすべてのゼロクロスを除外する。他の弁別器6およびしきい値発生器7による上記と同じ種類の二回目のフィルタ処理によって、最終的には、図2の(F)に例示されるようなデータセットが残る。前記弁別器6の出力に残った図2の(F)に例示したゼロクロスは、楽音の周期長さを定義する基本的なゼロクロスである。演算器8は、前記残ったゼロクロスのうち少なくとも2つのゼロクロスの間の時間間隔を検出し、その逆数を求める。該逆数は、その波形を分析すべき楽音の基本周波数に直接対応するものである。その結果としての周波数信号は、前記演算器8によって出力される音高信号に容易に変換されることができる。

【0036】本発明は、特に、爪弾きまたは打ち鳴らし操作によって励振される楽器（例えばギターのような弦楽器）から発生された音信号を入力してその音高を認識する場合において好適である。そのような場合、自然楽器から発生された該音信号から認識した音高情報を基にして様々な処理（例えば音高情報に基づく新たな楽音合成処理）を行なうように、利用できる。勿論、これに限らず、どのような音発生源から発生された音信号に対しても、本発明を適用することができる。

【発明の効果】以上のように、この発明は、簡単かつ確実に音高認識を行うことができる、という優れた効果を奏する。

【図面の簡単な説明】

【図1】ゼロクロスをもつ音信号の典型的な波形例を示す図。

【図2】この発明の一実施の形態に係る音高認識方法の処理手順を説明する図。

【図3】ゼロポイント近辺の信号波形の詳細を示す図。

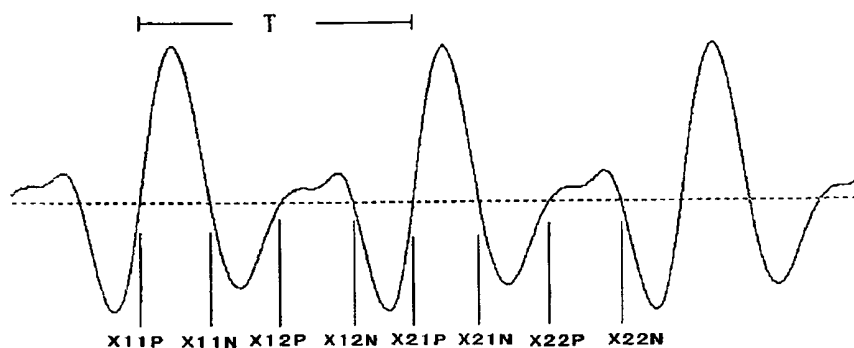
【図4】この発明の一実施の形態に係る音高認識装置を

示すブロック図。

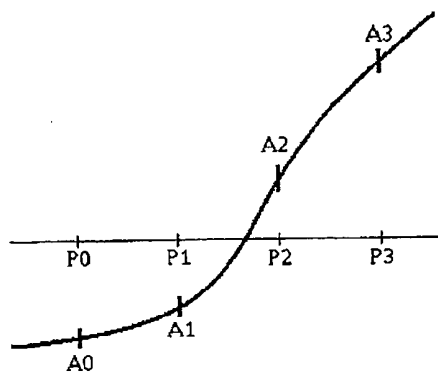
【符号の説明】

- T 周期長さ
- 2 ローパスフィルタ
- 3 a ゼロクロス検出器
- 3 b 急峻度演算器
- 4 弁別器
- 5 しきい値発生器
- 6 弁別器
- 7 しきい値発生器

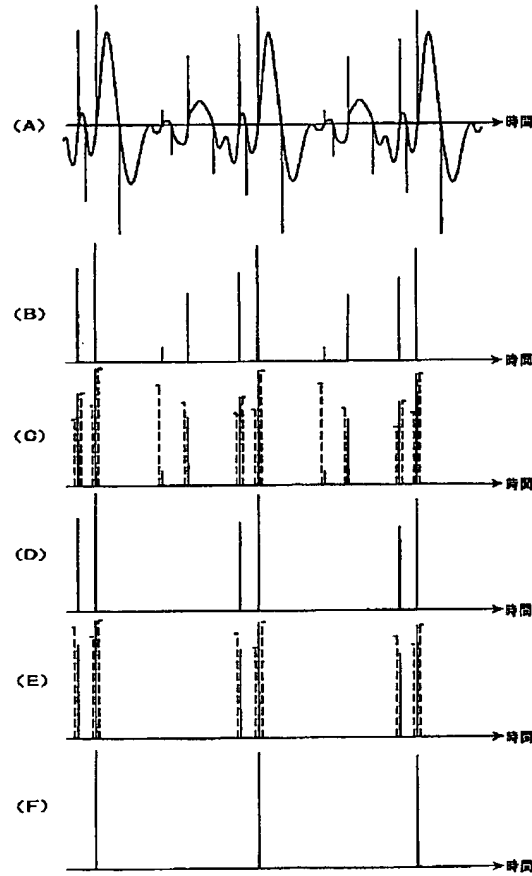
【図1】



【図3】



【図2】



【図4】

